

AD-A273 399

Reproduced From
Best Available Copy



IMAGE QUALITY ANALYSIS
OF COMPRESSED
SYNTHETIC APERTURE RADAR IMAGERY (U)

2

DTIC
ELECTE
NOV 30 1993

Todd D. Penrod

ADROIT SYSTEMS, INCORPORATED
FAIRBORN, OHIO

Gilbert G. Kuperman

CREW SYSTEMS DIRECTORATE
HUMAN ENGINEERING DIVISION
WRIGHT-PATTERSON AIR FORCE BASE OHIO 45433-7022

JANUARY 1993

FINAL REPORT FOR PERIOD OCTOBER 1991 TO SEPTEMBER 1992

Approved for public release; distribution is unlimited

93 11 29 023

AIR FORCE MATERIEL COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433

LABORATORY

93-29137

56px

NOTICES

When US Government drawings, specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise, as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Please do not request copies of this report from the Armstrong Laboratory. Additional copies may be purchased from:

National Technical Information Service
5285 Port Royal Road
Springfield, Virginia 22161

Federal Government agencies and their contractors registered with the Defense Technical Information Center should direct requests for copies of this report to:

Defense Technical Information Center
Cameron Station
Alexandria, Virginia 22314

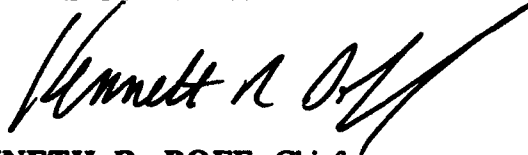
TECHNICAL REVIEW AND APPROVAL

AL/CF-TR-1993-0156

This report has been reviewed by the Office of Public Affairs (PA) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

FOR THE COMMANDER



KENNETH R. BOFF, Chief
Human Engineering Division
Armstrong Laboratory

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE Jan 1993	3. REPORT TYPE AND DATES COVERED Final Sep 91 - Dec 1992		
4. TITLE AND SUBTITLE Image Quality Analysis of Compressed Synthetic Aperture Radar Imagery (U)		5. FUNDING NUMBERS C: F33615-89-C-0532 PE - 62202F PR - ILIR TA - CH WU - 22		
6. AUTHOR(S) Gilbert G. Kuperman Todd D. Penrod, ADROIT Systems Inc.				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) ADROIT Systems Inc. 2970 Presidential Drive, Suite 340 Fairborn OH 45324		8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Armstrong Laboratory Crew Systems Directorate Human Engineering Division (HSC, AFMC) Wright-Patterson AFB OH 45433-7022		10. SPONSORING/MONITORING AGENCY REPORT NUMBER AL/CF-TR-1993-0156		
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) The study investigated the effects of image compression processing on synthetic aperture radar imagery. Image quality measurements were performed before and after compression. Image fidelity measurements were made comparing the compressed and original versions of each image. Operator performance was assessed using an interpretability rating scale. Possible effects of an interaction between compression and automatic target recognizer performance were explored using a generic constant false alarm rate automatic target cuer algorithm.				
14. SUBJECT TERMS synthetic aperture radar, image compression image quality, image fidelity, automatic target cuer			15. NUMBER OF PAGES 49	
			16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UNLIMITED	

THIS PAGE LEFT BLANK INTENTIONALLY.

PREFACE

This work was performed by ADROIT Systems, Incorporated, Dayton, Ohio, under subcontract 07-014-28S for Logicon Technical Services, Inc. (LTSI), Fairborn, Ohio, in support of the U.S. Air Force Armstrong Laboratory, Crew Station Integration Branch (AL/CFHI) under Contract Number F33615-89-C-0532. Mr. Robert Linhart was the Air Force Contract Monitor and Mr. Gilbert Kuperman (AL/CFHI) was the technical focal point for this effort.

The effort was supported by Armstrong Laboratory In-House Laboratory Independent Research (ILIR) funding as ILIR-92-9 and was performed under Work Unit ILIRCH22.

The purpose of the effort was to begin the investigation of the effects of data compression on Synthetic Aperture Radar (SAR) image quality. Digital image processing tools and methods were used to compress and measure/quantify the loss of SAR image quality and to begin to quantify the effects that this loss may have on human operator performance or on automatic target cueing systems. The study was conducted in the Visual Image Processing, Enhancement, and Reconstruction (VIPER) research and development facility of the AL.

Special thanks are due to Mr. Kenneth Crum (LTSI) who installed the image quality and image fidelity software modules on a graphics workstation and who was already ready to improve the utility of the measurement, analysis, and display interfaces. Thanks are also due to Mrs. Iris Davis (LTSI) who assisted during the image interpretability rating scale application portion of the study. The encouragement and suggestions of the attendees at several of the Automatic Target Recognizer Working Group (ATRWG) meetings held in 1991 and 1992 in the areas of clutter estimation are also greatly appreciated. Of course, the special talents and professionalism of the personnel of the Foreign Aerospace Science and Technology Center (FASTC), Wright-Patterson Air Force Base, Ohio, who served as subject matter experts in the operator performance portion of this research, must not go unrecognized. The cooperation and assistance of Mr. Robert Kaminski, who served as the point of contact at FASTC and who coordinated the participation of the imagery analysts, greatly contributed to the expeditious completion of that part of the study. The authors

owe a particular debt of gratitude to Ms. Denise Wilson (AL/CFHI) who provided guidance during the design of the experiment and who performed the statistical analyses reported herein; she shared freely of her experience in assessing the effects of compression of electro-optical sensor imagery on interpreter performance and served as a very knowledgeable peer during many of the phases of the study.

TABLE OF CONTENTS

<u>Section</u>	<u>Title</u>	<u>Page</u>
I	INTRODUCTION BACKGROUND	1
II	IMAGE COMPRESSION TECHNIQUES	2
	DATA COMPRESSION	2
	BACKGROUND AND RELATED STUDIES	2
	DISCRETE COSINE TRANSFORM COMPRESSION	4
	BLOCK TRUNCATION CODE COMPRESSION	11
	VECTOR QUANTIZATION COMPRESSION	13
III	IMAGE QUALITY MEASURES	18
	IMPULSE RESPONSE	18
	CONTRAST RATIO	20
	PEAK SIDE LOBE RESPONSE	23
IV	IMAGE FIDELITY MEASURES	24
	MEAN SQUARE ERROR	24
	SIGNAL-TO-NOISE RATIO	25
V	MEASUREMENT RESULTS	26
	EXPERIMENTAL METHODOLOGY	26
	IMPULSE RESPONSE MEASUREMENTS	27
	CONTRAST RATIO MEASUREMENTS	28
	MEAN SQUARE ERROR MEASUREMENTS	31
	SIGNAL-TO-NOISE RATIO MEASUREMENTS	32
	CORRELATION -- IMAGE FIDELITY	33
VI	OPERATOR PERFORMANCE	35
	SUBJECT MATTER EXPERTS	35

	TASK	35
	IMAGERY	37
	ANALYSIS AND RESULTS	37
VII	AUTOMATIC TARGET CUEING	39
	CLUTTER ESTIMATION	39
	RESULTS	41
VIII	CONCLUSIONS	42
	REFERENCES	43
	APPENDIX	46
	GLOSSARY	48

DTIC QUALITY INSPECTED 8

Accession For		
NTIS CRA&I	<input checked="" type="checkbox"/>	
DTIC TAB	<input type="checkbox"/>	
Unannounced	<input type="checkbox"/>	
Justification		
By		
Distribution /		
Availability Codes		
Dist	Avail and/or Special	
A-1		

LIST OF FIGURES

<u>Figure</u>	<u>Title</u>	<u>Page</u>
1	(a) Original signal $x(n)$ $0 < n < N-1$, (b) $2N$ point extension of $x(n)$, $y(n)$ $0 < n < 2N-1$.	6
2	(a) Even and odd members of $y(n)$ labeled for decomposition into $v(n)$ and $w(n)$, (b) $v(n)$, and (c) $w(n)$.	9
3	Example data set for Vector Quantization.	13
4	Scattergram of vector \mathbf{D} with code book vectors.	14
5	Block diagram of vector quantizer.	15
6	(a) Near ideal IPR, (b) Unweighted IPR, and (c) -35 dB Taylor weighted IPR.	19
7	Mean range IPR response.	29
8	Mean azimuth IPR response.	30
9	Mean Contrast Ratio measurements.	31
10	Mean Square Error Measurements of all Images.	32
11	Signal-to-Noise Measurements for all Images.	33
12	Response Control Panel Arrangement.	36
13	Mean Interpretability Ratings as a function of Image Compression Method.	38
14	Triple Window Filter.	39
15	Effects of Image Compression on TBIR-based Clutter Metric.	41

THIS PAGE LEFT BLANK INTENTIONALLY

Section I

INTRODUCTION

BACKGROUND

Current initiatives for the future application of high resolution Synthetic Aperture Radar (SAR) systems include distributed image processing, worldwide relays of unprocessed data and imagery, and multi-sensor tactical and strategic relays. Numerous advanced target acquisition concepts are based on the use of remote platforms to perform target search over large areas. Sensor imagery is processed, compressed, and transmitted to a strike vehicle where the imagery is interpreted either by a human operator, automatic target recognition (ATR) processing or a combination of the both (sometimes called ATC: automatic target cueing). This trend toward wider distribution and multi-sensor systems causes the consideration of significant bandwidth reduction, on the order of 10:1 or more.

A literature review revealed that while image quality metrics for electro-optical sensor systems are well established (less so for SAR), the image compression community applies "image fidelity" metrics in algorithm development and testing. No research was found which related image compression to degradation in image quality. Additionally, no research was found which related compression algorithm effects (compression induced image artifacts) to ATC performance. Sensor systems are typically tasked to search large areas for a variety of critical mobile targets and ATC systems are required to screen the volume of data for probable targets. High confidence operator confirmation is required to reduce the high false alarm rates associated with current and near-term ATC systems. Very limited operator performance literature was found which related image compression processing and target recognition performance. The purpose of this effort was to begin filling the void in research into the effects of compression on SAR imagery and its exploitation by either a human operator or an ATC.

Section II

IMAGE COMPRESSION TECHNIQUES

DATA COMPRESSION

In 1948, Claude Shannon [1] published his work defining the basis for calculating the information content of a data set and its theoretical minimum transmission rate. Since that time, various forms of quantization and data compression have been used in an attempt to meet and fall below Shannon's theoretical threshold. Data compression is possible primarily because of two basic observations. First, there is a large amount of redundancy in most images. That is, two picture elements (pixels) that are spatially close together tend to have similar magnitudes. Secondly, not all pixels carry the same information content. Just as in language, where some words may be irrelevant and can be omitted without causing a loss of understanding of the message. In imagery, a certain amount of pixel information may also be lost without loss in overall image quality. However, practical image compression processes are virtually always sub-optimal and result in some loss of image fidelity and image quality. Thus, the goal of image compression is to retain the best possible image quality for a given data rate.

BACKGROUND AND RELATED STUDIES

Compression techniques do not work equally well with all types of imagery. The spectral response of some recently developed sensor systems presents a challenge to data compression. SAR is perhaps the best example and presents the greatest challenges. SAR imagery is less literal than photographic and other forms of optical images. This is in part due to the operating frequency of the radar and in part due to coherent imaging and processing techniques. Typically, SAR systems use operating wavelengths around 1-5 cm as opposed to optical systems which use wavelengths around .4 - .7 μm . Thus, many objects which provide a diffuse scatter in visible light, appear smooth to the radar by providing a specular return. The cross section of these specular returns may change by several orders of magnitude as the viewing angle

changes by a few milliradians [2].

When closely spaced, randomly oriented reflectors are illuminated by a coherent source, the returned signal can constructively or destructively interfere, depending on the spacing of scatterers relative to the radar wavelength. These diffuse scatterers form a continuous background which is highly granular. This granularity is called speckle. The coherent imaging and processing used in SAR predisposes the imagery to speckle much like the speckle seen in coherent optical systems.

The combined effects of specularity and speckle in SAR images is that standard bandwidth compression techniques such as Differential Pulse Code Modulation (DPCM), which exploit data redundancy, do not work well with SAR images. Prediction of a sample value based on its neighbors cannot be easily met for SAR [2,3]. These same references argue that transform coding is not a viable alternative because the rapid decay of the autocorrelation function of the image requires large areas to be transformed. While a rather sophisticated transform method [4] has been suggested, the computational complexity encourages other methods to be considered.

Recently, several new image compression techniques have been developed which hold promise for compression of SAR images. It has been reported [2] that Block Truncation Coding (BTC) [5] has been used with moderate success in coding SAR (at a compression level of 2 bits/pixel). Vector Quantization (VQ) [6,7], which is sometimes referred to as a universal compression method [8], has been reported at a compression level of about 0.5 bits/pixel with optical images. It is anticipated that SAR imagery can be compressed to a more moderate rate.

Investigations to date on the use of compression with SAR imagery have been very limited. Only a few references to this topic were found and they all suffered from a common failing. That is, in comparing the decoded images with the original, only visual comparisons are made. In making these visual comparisons, none of the articles refer to the use of subjects which have been exposed to a large amount of SAR imagery or understand SAR image quality parameters in considering the visual effects of the compression/decoding process. Additionally, SAR systems being developed both within the DoD community and outside have very stringent image quality requirements

placed on them. These requirements are specific to SAR and SAR designers are typically required to develop detailed error budgets for contributors to image quality degradation. Compression is rarely, if ever, considered.

Three compression techniques (algorithms) have been developed for use in the present investigation. The first is Discrete Cosine Transform (DCT) compression. Despite its questionable performance [2,3] with SAR images it is the technique of choice in several image exploitation systems currently under development by the DoD. Additionally, a DCT hybrid compression is used in the Joint Photographic Exploitation Group (JPEG) image compression standard. The second method, BTC has been suggested as a good technique for use with SAR images [2]. Finally, VQ is considered. It is a newer method which has been used to compress electro-optical sensor images at greater than 10:1 compression ratios [9].

DISCRETE COSINE TRANSFORM COMPRESSION

DCT compression is a lossy transform technique. The DCT algorithm converts the spatial domain pixels of an image into cosine domain coefficients. The DCT is often used over other transforms because it provides good energy compaction, does not result in any blocking artifact at sub-image block edges [10] and results in a reduction of the correlation between coefficients in the new domain [11]. To accomplish the conversion, the image is first divided into sub-images. Each $N \times M$ sub-image is separately transformed into the cosine domain. The set of coefficients for each sub-image is then truncated, which equates to removing all frequencies higher than the coefficients which are kept. The remaining coefficients are then appropriately quantized and stored to represent the reduced bandwidth image. The stored coefficients can then be inverse transformed back to the image domain $N \times M$ sub-image.

In one dimension, the forward DCT is given by:

$$C(u) = \frac{2c(u)}{N} \sum_{n=0}^{N-1} f(n) \cos \pi \frac{(2n+1)u}{2N} \quad (1)$$

$$\begin{aligned} \text{where } & u = 0, 1, \dots, N-1 \\ \text{and } & c(u) = 1/(\sqrt{2}) \quad \text{for } u = 0 \end{aligned}$$

$$= 1 \quad \text{for } u = 1, 2, \dots, N-1$$

and in two dimensions is given by:

$$C(u, v) = \frac{4c(u, v)}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) \cos \pi \frac{(2m+1)u}{2M} \cos \pi \frac{(2n+1)v}{2N} \quad (2)$$

$$\text{where } u = 0, 1, \dots, M-1$$

$$v = 0, 1, \dots, N-1$$

$$c(u, v) = 1/2 \quad \text{for } u = v = 0;$$

$$= 1 \quad \text{for } u (v) = 1, 2, \dots, M-1 (N-1).$$

The original two dimensional data may be recovered by the inverse of the above:

$$f(m, n) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} c(u, v) C(u, v) \cos \pi \frac{(2m+1)u}{2M} \cos \pi \frac{(2n+1)v}{2N} \quad (3)$$

$$\text{for } m = 0, 1, \dots, M-1$$

$$n = 0, 1, \dots, N-1.$$

Direct implementation of (2) and (3) is computationally intensive due to the iterative nature of the algorithms. As is shown in Makoul [12], the transforms can also be realized using an $M \times N$ point two-dimensional Fast Fourier Transform (FFT).

Following Makoul's development of the fast cosine transform, he notes that (in one dimension) for a discrete time signal $x(n)$ the discrete Fourier transform $X(k)$ may be written as:

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-jnk\pi/N} \quad (4)$$

which is in general complex. Equation (4) can then be written in a general form as a sum of its real and imaginary parts [13]:

$$X(k) = X_r(k) + jX_i(k), \quad (5)$$

where $X_r(k)$ is the real part of $X(k)$ and $X_i(k)$ is the imaginary part of $X(k)$.

Using the definition of the discrete Fourier transform and adding and subtracting $\frac{e^{jnk\pi/N}}{2}$, one can expand equation (4) into its real and imaginary parts:

$$X(k) = \sum_{n=0}^{N-1} x(n) \cos 2\pi \frac{kn}{N} - j \sum_{n=0}^{N-1} x(n) \sin 2\pi \frac{kn}{N} \quad (6)$$

where

$$X_r(k) = \sum_{n=0}^{N-1} x(n) \cos 2\pi \frac{kn}{N} \quad (7)$$

and

$$X_i(k) = - \sum_{n=0}^{N-1} x(n) \sin 2\pi \frac{kn}{N}. \quad (8)$$

Using the above information as a basis for further development, Makoul [12] defines a new sequence $y(n)$ (Figure 1), to be the $2N$ even extension of $x(n)$:

$$y(n) = \begin{cases} x(n) & 0 \leq n \leq N-1 \\ x(2N-n-1) & N \leq n \leq 2N-1 \end{cases} \quad (9)$$

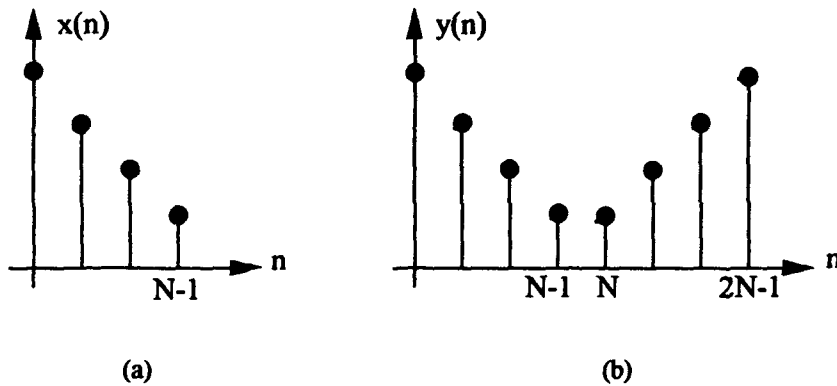


Figure 1. (a) Original signal $x(n)$ $0 < n < N-1$, (b) $2N$ point extension of $x(n)$, $y(n)$ $0 < n < 2N-1$

Once again starting with the definition of the FFT in (4) and applying it to the sequence $y(n)$, one obtains (using Cooley-Tukey notation):

$$Y(k) = \sum_{n=0}^{N-1} y(n) W_{2N}^{nk} \quad (10)$$

where $W_{2N} = e^{-j2\pi/2N}$.

Substituting (9) into (10), $Y(k)$ can now be written as:

$$Y(k) = \sum_{n=0}^{N-1} x(n) W_{2N}^{nk} + \sum_{n=N}^{2N-1} x(2N - n - 1) W_{2N}^{nk}. \quad (11)$$

Changing the variable of summation in the second addend so that it matches the summation of the first addend, one obtains:

$$Y(k) = \sum_{n=0}^{N-1} x(n) W_{2N}^{nk} + \sum_{n=0}^{N-1} x(n) W_{2N}^{2Nk} W_{2N}^{-nk} W_{2N}^{-k}. \quad (12)$$

Realizing that $W_{2N}^{2Nk} = 1$ for all integers k , factoring $W_{2N}^{-k/2}$ from each summation, and rearranging

$$Y(k) = W_{2N}^{-k/2} \sum_{n=0}^{N-1} x(n) [W_{2N}^{nk} W_{2N}^{k/2} + W_{2N}^{-nk} W_{2N}^{-k/2}]. \quad (13)$$

Remembering the exponential form of cosine is

$$\cos x = \frac{e^x + e^{-x}}{2} \quad (14)$$

the equivalent expression of (13) can be written as:

$$Y(k) = W_{2N}^{-k/2} \left[\sum_{n=0}^{N-1} x(n) \cos \pi \frac{(2n+1)k}{2N} \right]. \quad (15)$$

Comparing equation (1) with the bracketed expression in (15), one notes that with the exception of the amplitude weighting factor, $2c(u)/N$, they are the same. Therefore, it

is possible to write $Y(k)$ in terms of the Discrete Cosine Transform:

$$Y(k) = W_{2N}^{k/2} C(k) \quad (16a)$$

and conversely:

$$C(k) = W_{2N}^{k/2} Y(k). \quad (16b)$$

At this point, the DCT has been expressed in terms of a $2N$ point FFT. Other authors, [10,11,14] have provided similar treatment of the DCT. The method of Makoul [12] (to be developed next) provides a 2:1 (4:1 in the two dimensional case) reduction in computations by using decimation-in-time and an N point FFT.

Makoul (1980) defines two new sequences by dividing $y(n)$ such that:

$$\left. \begin{array}{l} v(n) = y(2n) \\ w(n) = y(2n+1) \end{array} \right\} \quad 0 \leq n \leq N-1. \quad (17)$$

Note that, in Figure 2, $v(n)$ contains all even numbered points ($n = 0, 2, 4, \dots$) of $y(n)$ and the other sequence, $w(n)$ contains all odd numbered points ($n = 1, 3, 5, \dots$) of $y(n)$. Further, $v(n)$ and $w(n)$ contain all of the points of the original sequence $x(n)$ and that the order of $v(n)$ data is reversed from the order of $w(n)$. That is,

$$w(n) = v(N-n-1). \quad (18)$$

Now substituting (17) into (10), the definition of the FFT of $Y(k)$, one obtains $Y(k)$ in terms of two N point FFTs:

$$Y(k) = \sum_{n=0}^{N-1} v(n) W_{2N}^{2nk} + \sum_{n=0}^{N-1} w(n) W_{2N}^{(2n+1)k}. \quad (19)$$

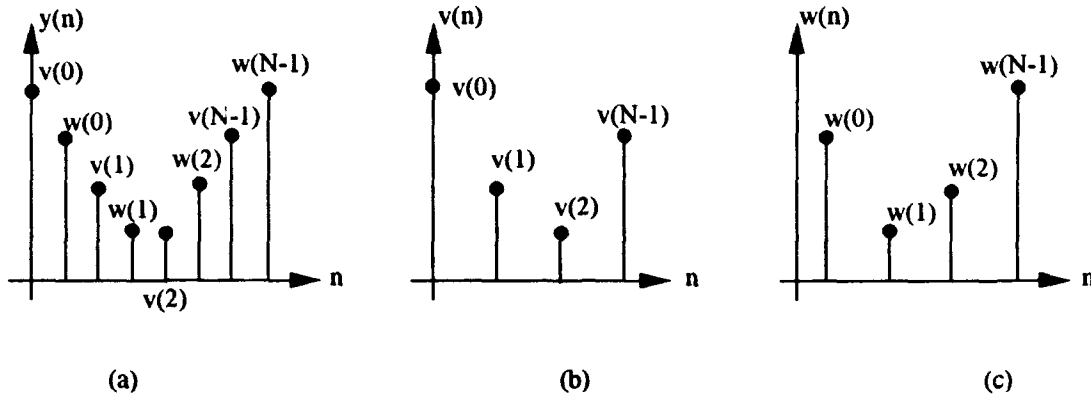


Figure 2. (a) Even and odd members of $y(n)$ labeled for decomposition into $v(n)$ and $w(n)$, (b) $v(n)$, and (c) $w(n)$

It is now possible to express (19) in terms of $v(n)$ by substituting (18) into the second addend:

$$Y(k) = \sum_{n=0}^{N-1} v(n) W_{2N}^{2nk} + \sum_{n=0}^{N-1} v(n) W_{2N}^{(2N-2n+1)k} \quad (20)$$

Following the development of equations (12) and (13), it is now possible to write

$$Y(k) = W_{2N}^{-k/2} \sum_{n=0}^{N-1} v(n) W_{2N}^{2nk} W_{2N}^{k/2} + \sum_{n=0}^{N-1} v(n) W_{2N}^{-2nk} W_{2N}^{-k/2} \quad (21)$$

Realizing that $W_{2N}^{2nk} = W_N^{nk}$ and that $W_{2N}^{k/2} = W_{4N}^k = W_N^{k/4}$, equation (21) can now be written as:

$$Y(k) = W_N^{-k/4} \sum_{n=0}^{N-1} v(n) W_N^{nk} W_N^{k/4} + \sum_{n=0}^{N-1} v(n) W_N^{-nk} W_N^{-k/4} \quad (22)$$

From the definition of the exponential form of the cosine function, equation (14), this becomes:

$$Y(k) = 2W_N^{-k/4} \left[\sum_{n=0}^{N-1} v(n) \cos \pi \frac{(2n+1)k}{N} \right]. \quad (23)$$

Since the cosine function has previously been shown in (6) and (7) to be the real part of an exponential function (4), equation (23) can be expressed as:

$$Y(k) = 2W_{4N}^{-k} \operatorname{Re} \left[\sum_{n=0}^{N-1} v(n) W_N^{nk+k/4} \right] \quad (24a)$$

or equivalently:

$$Y(k) = 2W_{4N}^{-k} \operatorname{Re} \left[W_{4N}^k \sum_{n=0}^{N-1} v(n) W_N^{nk} \right]. \quad (24b)$$

The bracketed portion of (24b) is simply the N point FFT of $v(n)$ multiplied by W_{4N}^k . The final step is to write the DCT of $x(n)$ as a function of $Y(k)$ as was previously shown in (16b):

$$C(k) = 2 \operatorname{Re} \left[W_{4N}^k \sum_{n=0}^{N-1} v(n) W_N^{nk} \right]. \quad (25)$$

Therefore, one sees that in one dimension it is possible to obtain the DCT of $x(k)$ using an appropriately decimated data set, $v(n)$, and an N point FFT. Using the separability of the cosine and Fourier transforms in the image domain, it is then possible to obtain the two dimensional DCT of image data using two one dimensional DCTs.

To implement the basic 2 bit per pixel DCT compression, each image was divided into multiple 8 x 8 non-overlapping sub-image blocks (i. e., $N = M = 8$). The mean of each sub-image was measured, removed from the block, and stored. The forward DCT was implemented using the fast DCT described above. The output of the transformed sub-image was an 8 x 8 set of cosine coefficients. Once all sub-images of the original image were transformed, the variance across each coefficient of all sub-images was determined (i. e., the variance of the (a,b) coefficient was determined by comparing the (a,b) coefficient of all sub-images). The variances were then ordered

from greatest to smallest. Only thirty coefficients with the greatest variances across the image were saved from each image. The thirty coefficients were then quantized to four bits using a Lloyd-Max quantizer [15,16]. Therefore, each 8 x 8 sub-image block was compressed 4:1; 64 picture elements (pixels) at 8 bits were compressed to 30 coefficients at 4 bits plus 8 bits for the sub-image mean.

Additionally, a small amount of overhead was required for each image. Thirty variances were stored with 7 bits (0-127) and the pointer to the position of the variance in the original data array was kept as a 4 x 4 (8 bit) pointer. These 450 bits were added to each image. For a 1K x 1K image, the overhead results in an insignificant 4.3×10^{-4} increase in bits per pixel.

BLOCK TRUNCATION CODE COMPRESSION

BTC is a lossy scalar compression technique which has been compared favorably with DCT and DPCM. The algorithm is designed to adapt to local statistical properties of small regions (blocks) of an image. As with DCT, the image is divided into $N \times M$ non-overlapping blocks. Local statistics (mean and variance) are then calculated for each of the blocks. Each pixel of the block is then coded with respect to the local mean using a two level (one bit) quantizer. To decompress, the number of pixels quantized greater than the mean is determined. An image block is then reconstructed using 2 levels (0-255) which are determined from the mean, the standard deviation, and number of pixels greater than (and less than) the mean.

The basic block truncation algorithm begins with the image block and calculates the first moment (mean), second moment, and sample variance of each block. The second moment is defined as:

$$\overline{X^2} = \frac{1}{N} \sum_{i=1}^N X_i^2 \quad (26)$$

The variance can then be determined as:

$$\sigma^2 = \overline{X^2} - \overline{X}^2, \quad (27)$$

where \overline{X} is the mean pixel intensity of the block.

Using the mean as the threshold for quantizing, code each pixel which is less than the threshold as "0" and each pixel which is greater than or equal to the threshold as "1." Thus the image block is represented by the mean, variance, and a bit plane of zeros and ones. In practice, the standard deviation is used more often than the variance because it is smaller and can be coded with fewer bits. The decoder must now take the information and reconstruct an image block. To accomplish this, the decoder determines q , the number of pixels which were coded as "1"s.

Remembering that the intent of the BTC is to preserve the first and second moments, one can write moment preserving equations [28] as:

$$n\bar{X} = (n - q)a + qb \quad (28a)$$

and

$$n\bar{X}^2 = (n - q)a^2 + qb^2, \quad (28b)$$

where n is the number of pixels in the block and a and b are the decoded output pixel values. Solving for a and b , it is now possible to write:

$$a = \bar{X} - \sigma \sqrt{\frac{q}{n - q}} \quad (29a)$$

$$b = \bar{X} - \sigma \sqrt{\frac{n - q}{q}}. \quad (29b)$$

The 2 bits per pixel compression using BTC was achieved using the algorithm discussed. As with DCT compression, the image was divided into blocks, 4×4 in this case. The mean and the standard deviation were calculated for each block and each was stored using 8 bits. The pixel intensity of the reconstructed image blocks was calculated based on equations (29a) and (29b). The intensities were then assigned according to the bit plane. Since all data for image reconstruction were stored in the mean, standard deviation and bit plane, no additional overhead was necessary. Therefore, the $4 \times 4 \times 8$ bit block (total 128 bits) was compressed to $8 + 8 + 16 \times 1$ bits (total 32 bits), exactly a 4:1 compression ratio.

VECTOR QUANTIZATION COMPRESSION

VQ is also a lossy image compression method. This technique is based upon the quantization of vectors (sub-image blocks) rather than scalars (individual pixels). It operates in the image domain and codes similar input vectors (similar means vectors which have approximately the same pixel intensities and order within the block) as one of a smaller set of vectors which is used to approximate the input vectors. Using this smaller set, called a code book or alphabet, VQ transforms the image vectors to minimum bandwidth channel symbols which are then stored. To decode, the channel symbols are matched to a look-up-table (LUT) to transform the symbols back to the appropriate alphabet letter (image domain blocks). The individual letters are then used as the image blocks to reconstruct the image.

As a graphical example of the process, consider the data set (image) with the data (pixel) values shown in Figure 3.

```
1 2 1 3 2 2 5 2 6
2 7 3 1 3 3 4 5 3 7
6 7 6 8 7 3 7 6 7 8
8 2 8 3 8 7 9 2 9 3
```

Figure 3. Example data set for Vector Quantization.

The data is then divided into a set of two dimensional vectors, $D = \{(1,2)(1,3)(2,2)(2,5)(2,6)(2,7)(3,1)(3,3)(4,5)(3,7)(6,7)(6,8)(7,3)(7,6)(7,8)(8,2)(8,3)(8,7)(9,2)(9,3)\}$. If one looks at the two dimensional scatter gram of the vectors as illustrated in Figure 4, it is possible to notice a natural clustering of the data points. It is then possible to estimate a suitable four letter code book to be, $C = \{(2,2)(3,6)(8,3)(7,7)\}$ by selecting points which are approximately in the center of each of the clusters. Each of the code book vectors can then be represented by channel codes 0,1,2,3, respectively. Thus the encoder would select channel code "0" to represent each vector in $D_0 = \{(1,2)(1,3)(2,2)(3,1)(3,3)\}$; each vector of the vector set $D_1 = \{(2,5)(2,6)(2,7)(3,7)(4,5)\}$ would be represented by channel code "1," and so on. At the decoder, a symbol "0" would be decoded as (2,2), a symbol "1" would be decoded

as (3,6). One could then draw a block diagram such as that in Figure 5.

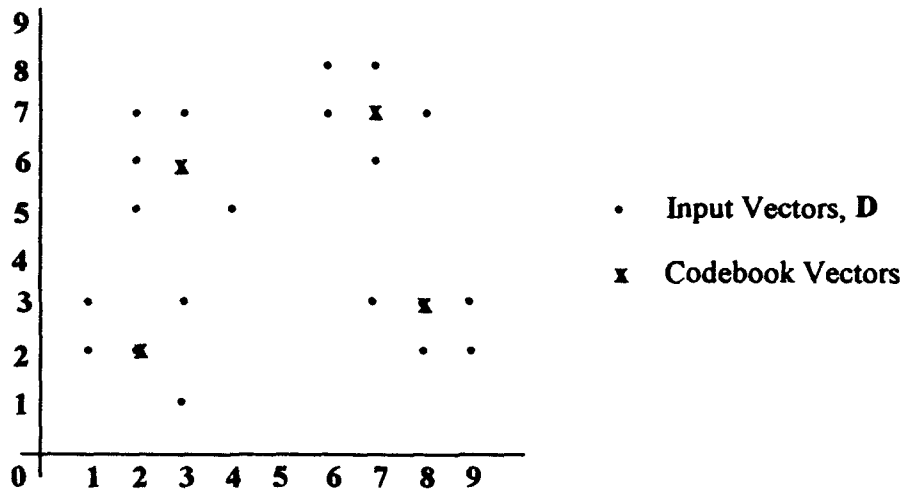


Figure 4. Scattergram of vector D with code book vectors.

While this visual method is suitable for small, two dimensional data sets, the difficulties in selecting appropriate vector neighborhoods grow rapidly as the vector dimensions and data set size increase and as the probability of any particular vector's occurrence takes on a unique probability function.

Therefore, before implementing vector quantization, it is necessary to derive an appropriate alphabet. As illustrated, a vector quantizing system operates by comparing each input vector with letters of the code book and storing (or transmitting) the channel symbol that corresponds to the "best" match between the two. The development of the best alphabet and the best match is determined in terms of the distortion measurement technique used. The distortion measure, $d(X, \bar{X})$ is an assessment of the cost for reproducing the input vector X by the alphabet vector \bar{X} . The minimum value of the distortion measure will identify the code word that is closest to the input vector.

Several distortion measurement techniques are typically discussed in the development of a code book. Among these are the Holder norm, Minkowski norm, quadratic distortion, and Itakura-Saito distortion measurement [6]. However, the most commonly used for image compression is the squared error [6,7], defined as:

$$d(X, \bar{X}) = |X - \bar{X}|^2 = \sum_{i=0}^{k-1} |x_i - \bar{x}_i|^2. \quad (35)$$

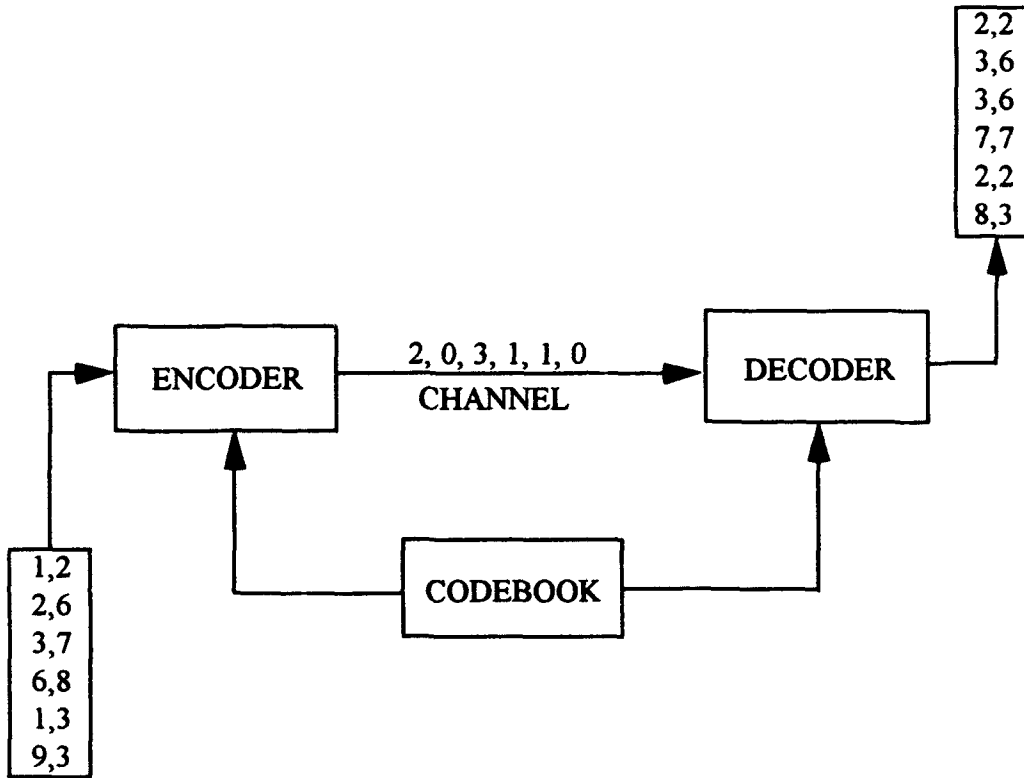


Figure 5. Block diagram of vector quantizer.

Mathematically, it is the simplest to implement and is well understood in that it represents the Euclidean distance between the two vectors X and \bar{X} . By taking the mean distortion over the image (adding the distortion of all vectors and dividing by the number of vectors), one then has the mean squared error.

The size of the code book is determined by the image fidelity and the data rate desired. As size of the code book increases, so does image fidelity. The data rate of a vector quantizer is given by:

$$R = \log_2 L \quad (36)$$

where R is in bits per pixel and L is the number of letters contained in the alphabet.

$$C_R = \frac{R}{MN}. \quad (37)$$

After choosing a method for measuring distortion, the alphabet can be calculated. Again several approaches to code book development are possible. One method is to assume an alphabet of the desired size. This can be done by taking the first 2^P vectors of the training sequence or by taking the same number of highly dispersed vectors from the data set. The initial code book is then used to cluster similar vectors (similar as determined by the chosen distortion measure) about each of the vector code words. The centroid of each cluster or neighborhood is then determined. The centroids then become the alphabet. This process is repeated until there is no (or little) change in the code book which result from consecutive iterations. Since the initial code book was randomly selected, the process is often called random code generation.

A second method for code book generation is known as splitting. This provides a means to grow a progressively larger code book from a one level quantizer. The centroid of the training sequence is typically used as the initial code vector. A second code vector is then chosen by either perturbing the first or by choosing the vector which is closest to the original vector. These two vectors are then used to separate the training vector into two groups, one coded by the original vector, and one coded by the perturbed vector. The centroid of each neighborhood is then calculated and the two code vectors (the centroids) are again perturbed to form four vector code words. The entire process is repeated until a code book of the desired size has been developed.

The foregoing clustering techniques, which require an initial code book and then iteratively improve it, are widely known as the LBG algorithms after Yosef Linde, Andre Buzo, and Robert Gray who first proposed the methods for vector quantization [7]. While computationally intensive, they provide a systematic means for finding neighborhood patterns and therefore, allows optimization of the code book. The methods are, however, an extension of Lloyd's algorithm [15] developed for scalar quantization.

To implement the two bit per pixel VQ compression for this study, a small

training image was used for code vector development. The training image was divided into 2×2 vectors and a 256 letter alphabet was created. The distortion measure used for the alphabet development was the squared error. The alphabet was grown from the mean of the training image by splitting. Once the alphabet was developed, the input images were divided into 2×2 vectors and matched to the code vectors using the minimum squared error criteria. The output channel code was quantized using 8 bits (256 levels). Therefore, the $2 \times 2 \times 8$ bit input vectors (total 32 bits) were compressed to a single 8 bit channel code, providing an exact 4:1 compression.

Section III

IMAGE QUALITY MEASURES

In order to adequately discuss image compression, it is necessary to develop measures of effectiveness for evaluation of the compression techniques. While compression ratio (ratio of input bits/pixel to output bits/pixel) is an important measure, it is not totally adequate in that it does not provide a measure of change in image fidelity when compared to the original image or a loss of original image quality. Therefore, one must develop the means to measure the change in information after compression. For this study, a number of classical SAR image quality measures which are typically used in the specification of SAR systems were integrated into a utility to study image quality.

IMPULSE RESPONSE

The classical performance measure used to rate most imaging systems is resolution, or their ability to provide separate images of closely spaced objects. In general, for optical systems, measured resolution depends on physical size of the measured object and contrast of the scene. Since SAR depends upon coherent phase detection of the reflected microwave energy, another dimension is added to characterization of its performance. When using coherent imaging and processing, the reflected signals from scatterers in a complex object may constructively or destructively combine depending on the phase relationship (geometry of the scatterer). This will enhance or eliminate amplitude nulls between scatterers in the objects. This means that the measured resolution of a SAR can change significantly as the precise wavelength spacing of the imaged objects change in half wavelength multiples.

As a means to eliminate the apparent change in resolution caused by geometry dependent phase of the SAR, a linear systems approach is used to provide measure of system resolution. Linear Systems Theory often characterizes linear systems by their impulse response. The SAR image formation process is a linear system which uses

pulse compression to obtain range resolution and Doppler processing to obtain cross-range (azimuth) resolution. It will, therefore, have a two dimensional intensity valued impulse response.

The definition of the SAR IPR measurement then, is the two dimensional measured or calculated image domain (spatial domain) response of the SAR system to a point target input. It is most typically measured along the output sampling axis (range and azimuth). IPR width is usually calculated at the -3 dB (half power) point but, on occasion, the -6 dB or -15 dB widths may also be of interest as indicative of the separability of two weak signals in proximity to each other.

The desired impulse response function would be a very narrow spike with no sidelobes (Figure 1a). In practice, this cannot be achieved and a typical imaging radar impulse response (unweighted) will look like Figure 1b. The IPR performance is then modified by the effective system transfer function (which includes radar and processor). For many SAR systems, a -35 dB Taylor weighting is applied to achieve an impulse response as shown in Figure 6c.

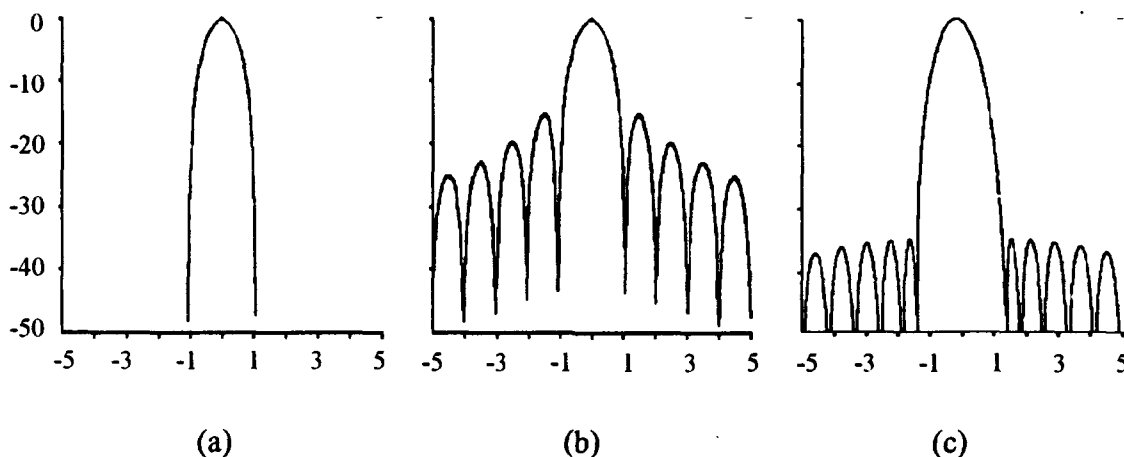


Figure 6. (a) Near ideal IPR, (b) Unweighted IPR, and (c) -35 dB Taylor weighted

For this study, the -3 dB width in range and azimuth (cross range) will be used. The operator selects the point return to be measured. The peak and the two pixels immediately to each side in the x direction are selected. A similar selection is made in

the y direction. The peak pixel, along with the selected adjacent pixels define a quadrant. The pixels within the quadrant are interpolated using a gaussian interpolator at one-half the distance to form a 9 x 9 pixel array. The array is searched for a maximum. The maximum level and its location are stored as a refined peak location. The -3 dB level is calculated from the peak of the interpolated data . The interpolated values obtained above are searched on both sides of the peak for the pixel which straddles the -3 dB level. Linear interpolation is then used to find the -3 dB width. Since the images are from various sources, no scaling of pixels to ground distance is done.

CONTRAST RATIO

Contrast ratio (CR) is perhaps the most general and all encompassing description of SAR image quality. It provides a measure of the effects on the final image of system power, antenna gains and losses, and noise from various sources throughout the radar system and image formation process. A high CR produces a crisp, distinct image while a low CR produces a "washed out" image with little clarity. From a visual perspective, the basic idea of SAR CR is roughly analogous to the CR in optical images, but they are not directly comparable. SAR image CR is a somewhat loosely defined term which is used to measure the ratio of the signal plus total noise to the total noise present in an image. In general, noise in the image is attributed to two noise source components, one signal dependent and the other signal independent.

Signal dependent noise will contribute an amount of noise proportional to the average image intensity. This type of noise generally results in additional side lobe energy from strong target returns and is represented by the term ISL (Integrated Side Lobe level). ISL represents all energy returned from a scatterer not contained in the main lobe of the processed data. Figure 6b in the previous section illustrates a typical system response to a point target. It is composed of the center or main lobe and a side lobe structure. The energy in the side lobes taken as a whole is characterized by the term ISL. ISL can be the manifestation of sinusoidal non random phase and amplitude errors (nonlinearities) in the signal processing chain which can cause several image quality degradation anomalies. Strong "spikes" near a strong return can result in false target detection. In some cases, weak returns may be swamped by nearby large

average backgrounds or large, strongly reflecting targets resulting in some targets going undetected. Random nonlinearities in the signal processing chain appear as uniform noise which causes a "fill in" or raising of the overall sidelobe level which may decrease visible separation (resolution) of closely spaced objects.

Signal independent noise sources contribute a fixed amount of noise regardless of image amplitude. The primary contributor here is receiver thermal noise, N_t . Other sources such as bit errors and digital round off errors can also add significantly to signal independent noise and are generally considered as additive noise, N_A . This signal independent noise acts to reduce contrast by adding random elements to the scene, producing imagery that resembles a "snowy" television picture. This noise can be overcome by increased antenna gain, transmitter power, and lower system losses, however, aircraft physical and environmental constraints limit what is practically achievable.

Thus, poor image CR results in a "washout" of the image. This washout can include *de emphasis of boundaries, fill in of no- and low-return areas, spikes, and "snowy" images*, all of which reduce the utility and exploitability of the image.

In the image domain, the measurement of total noise (no return area) must be done in a portion of the image where "no signal" exists. Such no-return areas may be paved roads, smooth surfaces or very smooth water where almost no signal is returned to the radar. The measured residual noise in the no-return area is an approximation of the total system noise, S_N . S_N is then defined as:

$$S_N = N_t + ISL + N_A \quad (38)$$

where ISL is the Integrated Sidelobe Level, a measure of the contribution of sidelobe energy from surrounding background. The number and size of the no-return area will have an influence on the stability and accuracy of the measurement. Measuring a large number of no return areas provides a stability to the measurement. The size of the no-return area will affect accuracy as well; too large an area and near-in sidelobe effects are diminished (ISL not accurately measured, and not fully accounted for), too small and the chances of mainlobe clutter interference (inclusion of some non-noise signal, C ,

as described in equation 39) are more likely resulting in an inaccurate CR measurement.

Signal measurement is performed in an average background. The measurement of the average background signal presumes a rather distributed background clutter level which is not commonly found in nature. The measured signal return from this area is:

$$S_c = C + N_t + ISL + N_A \quad (39)$$

where C is the average signal level due to clutter.

CR is defined as the ratio of the average intensity of a distributed clutter background to the average intensity of a no-return area and has the units of decibels. That is:

$$CR = 10 \log_{10} \frac{\text{Signal} + \text{Total Noise}}{\text{Total Noise}} = 10 \log_{10} \frac{C + N_t + ISL + N_A}{N_t + ISL + N_A} \quad (40)$$

The "looseness" in the CR definition comes from the difficulty in obtaining controlled measurement conditions. That is, the size of the no-return areas and the definition of average background are not always clear. From a user's perspective, the measurements are not terrain to low return area but rather "grass to shadow" measurements or "trees to shadow" measurements. Typical values for CR are 1 dB to greater than 15 dB; however, 8 dB is usually considered minimally acceptable.

One final note: the formulas previously described assume that at the output of the processor, the In-phase and Quadrature-phase (I and Q) channels are detected using the sum of the squares. That is, the pixel value is calculated using:

$$\text{Pixel Value} = I^2 + Q^2 \quad (41)$$

or some reasonable approximation of the formula which provides a power representation of the processed signal. If on the other hand, the output of the processor is detected using an energy representation such as the square root of the sum of the

squares:

$$\text{Pixel Value} = \sqrt{I^2 + Q^2}, \quad (42)$$

then CR is correctly calculated:

$$\text{CR} = 20 \log_{10} \frac{\text{Signal} + \text{Total Noise}}{\text{Total Noise}} \quad (43)$$

PEAK SIDE LOBE RESPONSE

The purpose of the Peak Sidelobe Response is to measure the dispersion of energy outside of the mainlobe and in the primary sidelobe of the system impulse response. Sidelobes are a result of the finite processing aperture and finite filter lengths. They can be controlled using weightings but weightings increase mainlobe width and won't remove the affects of errors in the data, another potential causes of sidelobe problems. Sidelobes due to the processing aperture may stay at a constant level for the near-in sidelobe region (e. g., Taylor weighting), but will fall off at some predetermined rate away from the interpolated peak. Low frequency deterministic phase errors will raise near-in sidelobes and high-frequency deterministic phase errors will raise sidelobes farther out. Narrow-band random phase errors will raise an ensemble of sidelobes (e. g., a pedestal effect). Wideband random phase errors tend to raise all sidelobes. The 2-D definition is again aimed at the multiple sidelobe axes that result from separate processing in range and azimuth.

Peak sidelobe response was not used in the study of effects of compression on image quality. It is briefly discussed here because it is a commonly SAR image quality metric. The IPR metric, which was applied to the study imagery, is somewhat related.

Section IV

IMAGE FIDELITY MEASURES

The objective of image compression is to be able to store or transmit a coded image in as few bits as possible, while maintaining the information content of the original image. Image fidelity measures are used to characterize the departure of the compressed and recoded image from the standard or original image. Thus image fidelity becomes an important measure of the "goodness" of a compression method.

The downfall of most image fidelity measures is that they measure absolute change in the image as opposed to information content or intelligibility. Quantitative measures aimed at both are often computationally cumbersome and inaccurate outside of a small well defined image set.

Since no measures for image fidelity have been developed that are well understood, work well across all image sets, and relate to information content, two of the more classical image fidelity measures were chosen for this study. The first is Mean Square Error (MSE) and the second is Signal-to-Noise Ratio (SNR).

MEAN SQUARE ERROR (MSE)

The MSE provides a pixel by pixel measure of the image change from the pre-compression to the post compression image. It is the most widely used image fidelity measure in compression studies in open literature. The popularity of the measure undoubtedly stems from its simplicity (mathematically) and ease of implementation. The MSE is defined as:

$$MSE = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} [g(m,n) - f(m,n)]^2 \quad (44)$$

where M and N are the number of rows and columns of the images and f(m,n) and g(m,n) are the (m,n) pixels of the original and compressed image, respectively. Note

that this measurement does not differentiate between a few large errors and many small ones. Further, in certain instances there appears to be little correlation between this pixel by pixel measure and the visually perceived information content.

SIGNAL-TO-NOISE RATIO (SNR)

SNR is a more global (less pixel oriented) measure of the image change than MSE. It is less widely used than MSE. While it is computationally simple and mathematically well understood, it has not shown significant correlation with operator performance or information loss. If the difference between the output (compressed and recoded) image and the input image is considered to be "noise," so that each output pixel consists of an input signal plus noise (the error); then the output image pixel is defined as:

$$g(m,n) = f(m,n) + e(m,n). \quad (45)$$

SNR is defined as:

$$SNR = \frac{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} [g(m,n)]^2}{\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} [g(m,n) - f(m,n)]^2}. \quad (46)$$

Note that SNR as it is calculated here is dimensionless. For this reason, it is sometimes expressed in dBs. However, for this study, it is left as a dimensionless figure of merit.

Section V

MEASUREMENT RESULTS

EXPERIMENTAL METHODOLOGY

Four area coverage SAR scenes from a 1982 Operational Utility Evaluation (OUE) of a developmental SAR sensor system were provided by the Armstrong Laboratory for evaluation. Scenes were of Barstow CA., Castle AFB CA, Vandenburg AFB CA, and the Long Beach Naval Shipyard. Scene content included:

- Railhead, convoy, howitzers, tanks, USMC Supply Center, roads, self-propelled artillery
- Airfield, B-52s, KC-135s
- Shipyard, storage area, ships
- Missile launch and military facilities

Each of the scenes was divided into sixteen (16), 512 x 512 pixel subimages to provide an image of usable size and a reasonable size sample set. From the resulting set of 64 images, 23 were chosen to be studied based on image content, number of appropriate point like returns, and areas for CR measurements. The selected images represented 8 of the 17 STANAG target categories. Each image was subjected to three different compression algorithms: DCT, BTC, and VQ, as previously described. Once compressed, the image files were expanded back (i. e., decompressed or reconstructed) to their original size of 8 bits per pixel. The final experimental set of images, therefore, consisted of 92 images (4 x 23).

Image quality was determined for each of the 92 images. CR was measured four to six times for each image, in at least two different areas of the image. Exact pixel coordinates of the vertices of the polygon were recorded so that CR for all four images would be made over the same area. The CR measurements for each image were

averaged and compared to the average CR of the original image. Impulse response measurements were made on between 2 and 4 point-like returns within each image. Again coordinates were recorded so that the same points could be measured in each of the image.

Image fidelity measures were made for each of the compressed images using MSE and SNR as previously described. These measurements quantify the overall change in the resultant (i. e., post-compression) image as compared to the original and are computed across the entire image; therefore, only one measurement was made for each of the compressed images.

IMPULSE RESPONSE MEASUREMENTS

Ideally, IPR measurements should be taken from calibrated point targets. Due to the origin of the images, limited numbers of known point targets were available. IPR measurements were taken on points that were visually estimated to approximate a point target in the original image. The results of the -3 dB IPR measurements are summarized in Figures 7 and 8 which provide the normalized range and azimuth IPR measurement results separately. In general, compression causes an increase in the IPR response. Note that it appears that DCT most often preserves the original IPR response the best, and over all images, degrades the IPR the least. In a few isolated cases, DCT even appears to enhance IPR response.

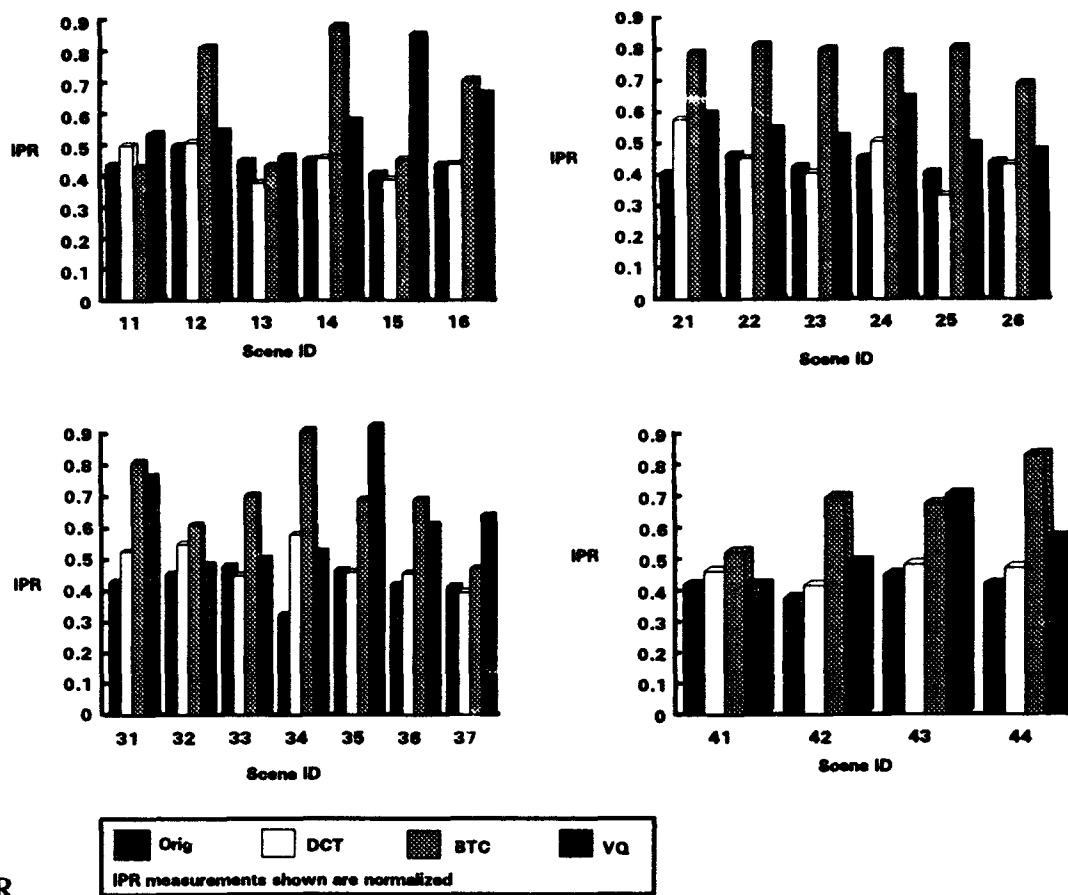
A paired-comparison t-Test was performed to investigate the effect of the three image compression techniques on the IPR measure of image quality. In applying the t-Test, all pairs of original and compressed images were compared for each compression technique. Statistically significant changes (both $p < 0.01$) were found for both the BTC and VQ techniques. The greatest degradation in image quality was found for the BTC compress and the IPR measurement made in the range direction. The mean loss in IPR was 62 percent for this combination.

CONTRAST RATIO MEASUREMENTS

As previously discussed, SAR image CR is a measure of the signal plus noise to the total noise present in the image. Ideally, these measurements are taken from areas with controlled no-return pads and known high return areas. This is done so that CR between images can be easily compared and to help control the amount of near-in sidelobe and mainlobe clutter that affects the measurements. Since controlled no-return areas were not always present in the images, low return areas such as shadows, roads or runways were used. These were sometimes smaller than ideal, and caused some less than optimal results. This is potentially the cause of some of the variance in the observed trends in the contrast measurements. Figure 9 provides the results of the CR measurements.

In general, it appears that BTC causes the smallest changes in CR (BTC results in the smallest CR change in over half of the images). This can be explained by looking at the way BTC compresses. BTC uses a very small local area (a 4 x 4 block) of the image and preserves the mean pixel amplitude of the area. CR is a measure of the ratio of the pixel amplitude in a high and a low return area. Since the high and the low return areas in the compressed image should contain several blocks which have the same average value as in the original image, minimal change on CR should be expected.

The paired-comparison t-Test procedure was also applied to the CR measurements. No statistically significant difference was found between the original and compressed imagery sets.



IPR

Figure 7. Mean range IPR response

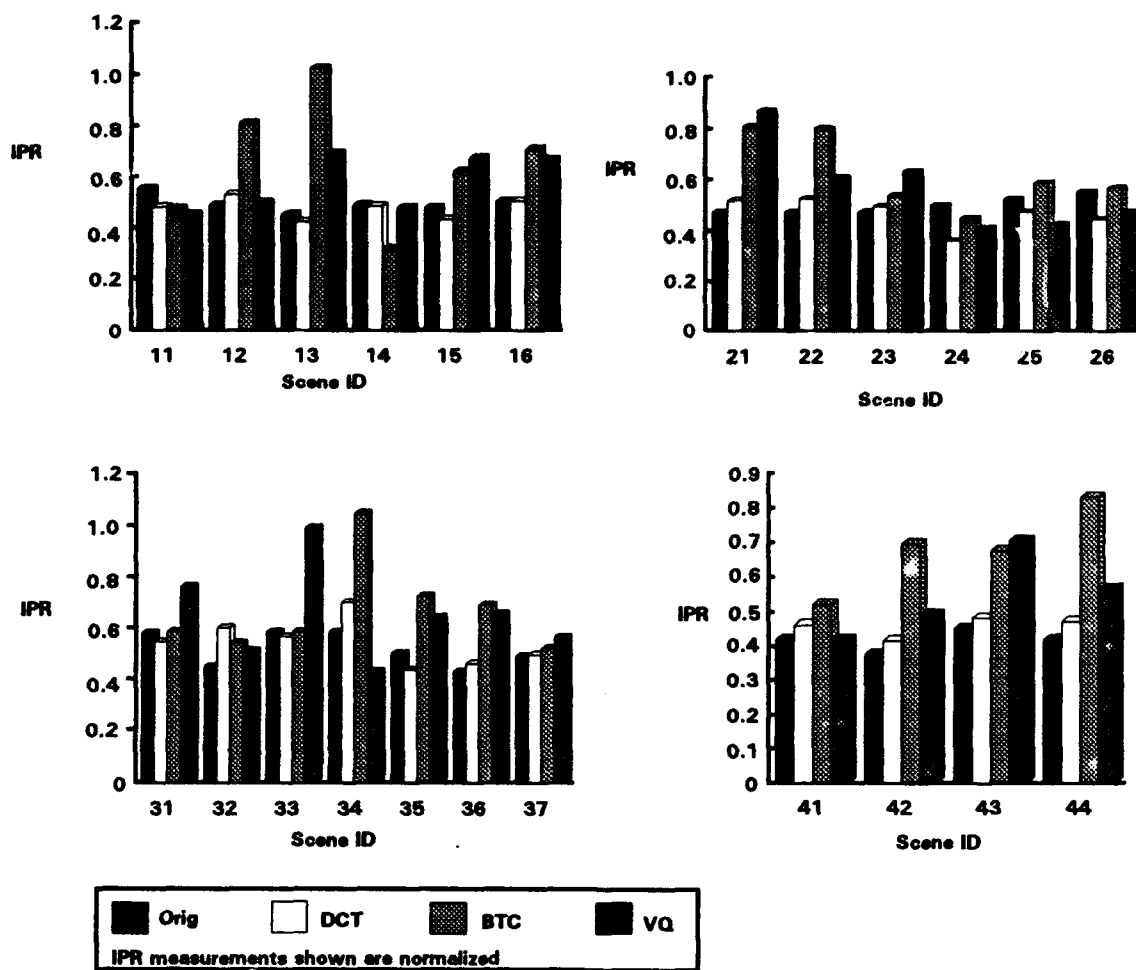


Figure 8. Mean azimuth IPR response.

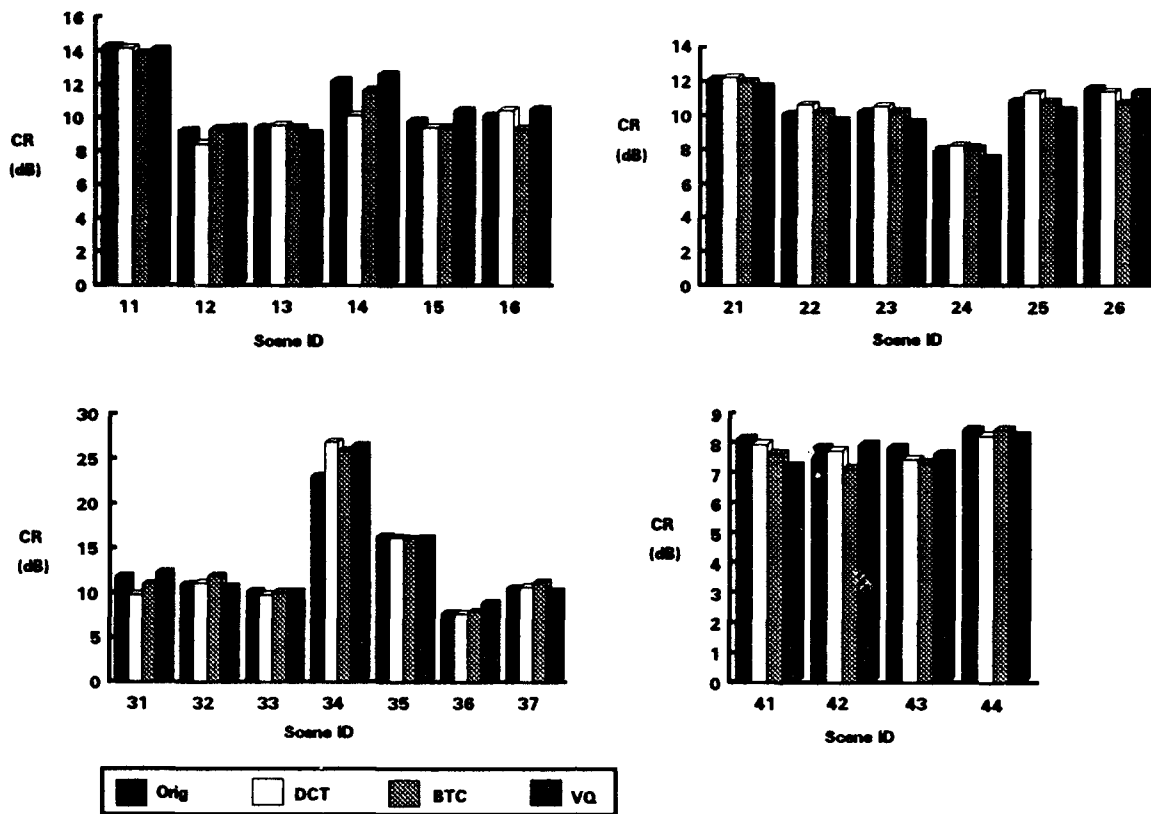


Figure 9. Mean Contrast Ratio measurements.

MEAN SQUARE ERROR MEASUREMENTS

MSE was calculated across each of the images. Results are provided in Figure 10. Notice, that at least for the specific implementation of VQ used in this study, it consistently outperforms the implementation of both DCT and BTC. This is somewhat to be expected because as VQ has been implemented here, the coded vector selection is based on minimum MSE. It is important to remember, though, that MSE by itself does not indicate if the overall error is due to a large number of small errors or a few very large errors. Also, MSE has not been shown to correlate with information content or interpretability.

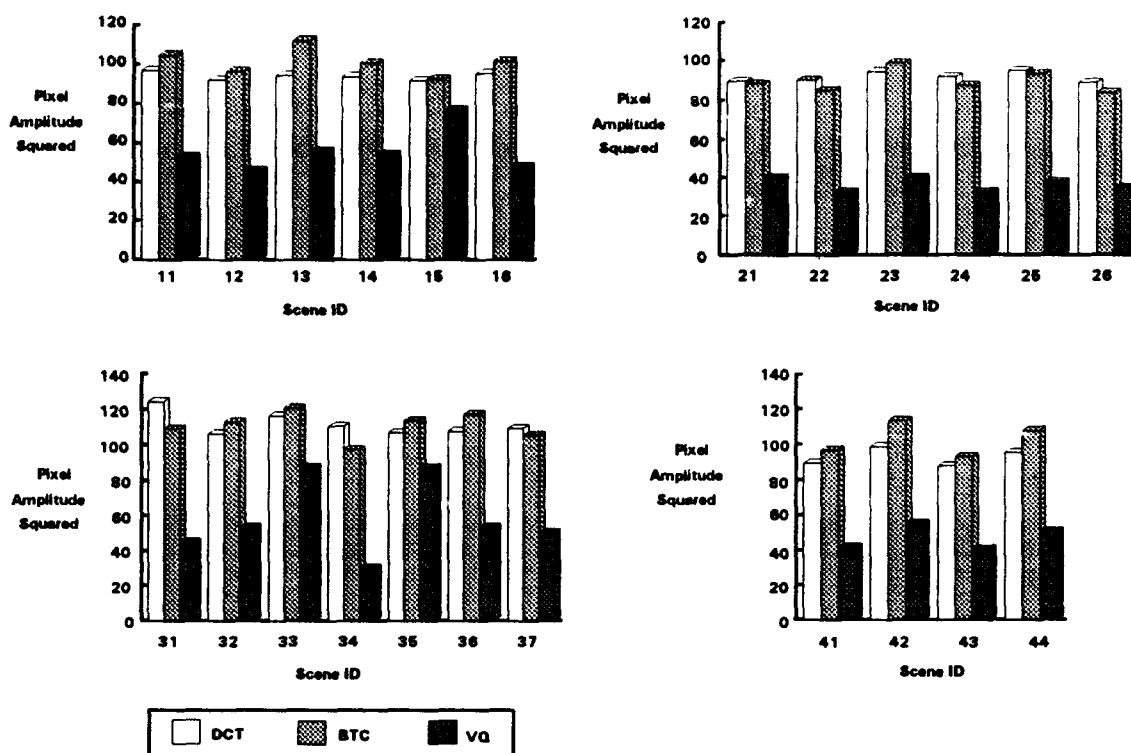


Figure 10. Mean Square Error Measurements of all Images.

SIGNAL-TO-NOISE RATIO MEASUREMENTS

Figure 11 shows the results of the SNR measurements. As with MSE, VQ again outperforms the other compression techniques. If one examines the formulas for these calculations, one notices that the denominator of the SNR (or the error component) is the squared error as determined in the MSE (before dividing by the size of the image). It follows then that a minimum MSE will be associated a maximum SNR.

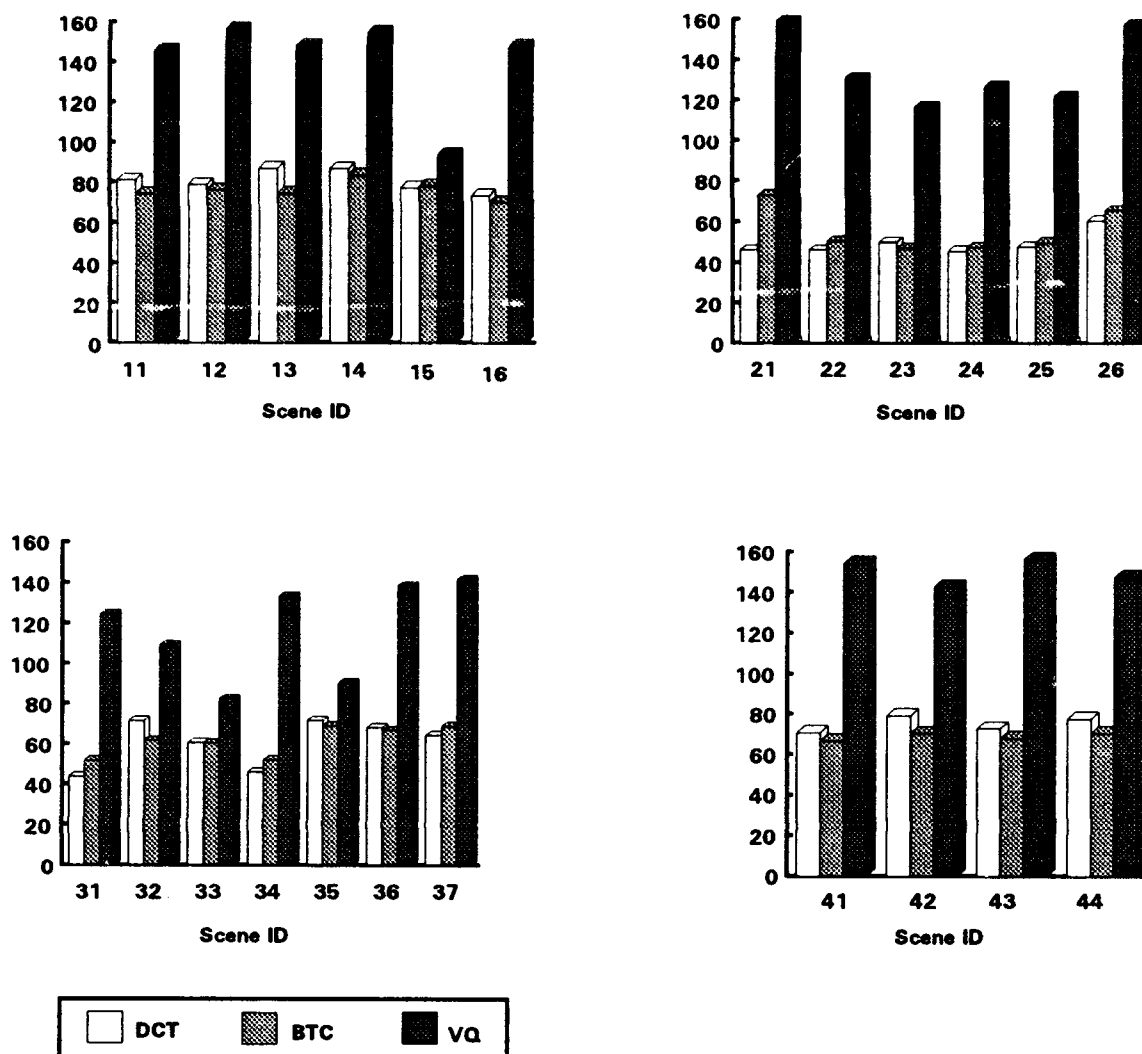


Figure 11. Signal-to-Noise Measurements for all Images.

CORRELATION--IMAGE FIDELITY

Since both MSE and SNR image fidelity metrics were computed, it was of interest to compare the two measures. Pearson's correlation coefficient, P_{CC} , was computed for each of the image compression techniques. Statistically valid agreement between MSE and SNR was found only for the VQ compression technique ($P_{CC} = -0.65$, $p < 0.01$).

It was also of interest to investigate the relationship between the image fidelity

measures (MSE and SNR) and the image quality measures (IPR and CR). The P_{CC} was computed for all paired imagery and for each image fidelity and image quality metric. No significant correlation was found between either image fidelity metric and CR. A significant P_{CC} ($P_{CC} = -0.44$, $p < 0.05$) was, however, found between MSE and IPR (measured in the azimuth direction) for the VQ compression technique. With regard to range IPR, both the BTC and VQ techniques yielded significant P_{CC} ($P_{CC} = 0.50$, $p < 0.05$, for BTC; $P_{CC} = 0.43$, $p < 0.05$, for VQ).

Section VI

OPERATOR PERFORMANCE

SUBJECT MATTER EXPERTS

Twelve imagery analysts (IAs) assigned to the the Foreign Aerospace Science and Technology Center (FASTC), Wright-Patterson Air Force Base, Ohio, served as subject matter experts during this portion of the study. Ten of the IAs were male (nine civilian and one military) and two were female (both civilian). All of them were experienced in image analysis and exploitation. Their mean experience level was 5.6 years. Two of the IAs had received both basic and advanced imagery interpretation training at military service schools in addition to on-the-job training. All IAs reported working with SAR imagery in their current assignment. Four reported exploiting SAR on a daily basis, three reported frequent exploitation of SAR, while the remaining five reported working with this medium on an infrequent basis. All IAs reported having received specific training on the use of image interpretability rating scales (described below) and all IAs reported using such scales as part of their current duty assignment.

TASK

An image interpretability rating scale was employed. There were ten rating categories, graduated as zero through nine, available for use in reporting. Each category was accompanied by semantic anchors to support response consistency. The semantic anchors described specific examples of information extraction associated with ground, air, and naval order of battle reporting categories. In the rating scale, a rating of 0 meant that the image was useless for exploitation purposes. Levels 1 through 3 were associated with detection and object counting tasks, level 4 and 5 referenced target classification tasks, and level 6 through 9 provided examples of target identification tasks. In imagery interpretation, detection refers to discovering the existence of one or more objects (i. e., an object of possible military interest is present in the image). Recognition or classification, a higher level of information extraction, is

associated with performing discrimination between target classes (i. e., the object is a tank [rather than a truck]). Identification is associated with performing discriminations within a single target class (i. e., the object is an M-1 main battle tank).

The analyst was seated comfortably in a reduced illumination booth. A 13 inch (33 cm) diagonal, monochrome (P47, green phosphor) ELECTROHOME Model 38-V19NDA-BP monitor. The booth and reduced illumination protected the subject from extraneous glare, distracting noises, etc. A special purpose pushbutton response panel was located on a pull-out work surface located in front of the monitor. Figure 12 depicts the arrangement of the response buttons.

0	1	2	3	ENTER
	4	5	6	
	7	8	9	READY

Figure 12. Response Control Panel Arrangement

The brightness and contrast controls of the monitor were adjusted by the experimenter with reference to a ten step linear intensity gray scale image display. The adjustment was such that each step in the gray scale was clearly discernible, without overdriving the monitor. The controls were then "locked."

When the subject was ready for the experiment, a medium gray field (intensity value 129) containing a centered fixation "+" was displayed on the monitor. The subject controlled the timing of imagery presentation. When the subject was ready to rate an image, the READY button was pushed. The subject indicated the appropriate rating for that image by pushing one of the buttons marked "0" through "9." The subject could change the rating simply by depressing a different button. While the subject was participating in the experiment, a paper listing the rating category definitions (semantic anchors) was available for reference. No time limit was imposed on the task. When the subject was satisfied with the rating assigned to an individual image (version), the ENTER button was pushed and the (last) rating (entered) was recorded. The subject repeated the READY-rating-ENTER sequence until all images had been rated.

IMAGERY

Each image was presented in four different versions: original (i. e., uncompressed) and after DCT, BTC, and VQ compression and reconstruction. It was felt that the subjects were experienced in making independent image interpretability rating scale judgements and would not be unduly biased by the repetition of image content. Each subject, however, did receive the images in a unique random order. That is, the original, DCT, BTC, and VQ versions of each scene, as well as the scenes themselves, were presented in a distinct randomization to each subject.

ANALYSIS AND RESULTS

A statistically significant difference (ANOVA, $p < 0.01$) was found between the three compression methods and the original (uncompressed) imagery set. A *post hoc* statistical test (Tukey's Honestly Significant Difference) revealed that the difference was entirely due to the comparison between the original and BTC-transformed image versions ($p < 0.05$). That is, the original images and the DCT and VQ versions formed one group which was significantly different from the DCT, VQ, and BTC versions, also considered as a group.

The mean interpretability ratings, pooled across IAs and images, are depicted in Figure 13. As can be seen in the Figure, each of the image compression algorithms introduced a slight loss in image utility for exploitation (as quantified by the interpretability ratings) when compared against the original version of the images. The mean rating for all original images and all IAs was 2.83. The BTC algorithm introduced the greatest degradation in image interpretability, reducing the mean rating to 2.52. (DCT supported a mean interpretability rating of 2.69 and VQ a mean rating of 2.71.)

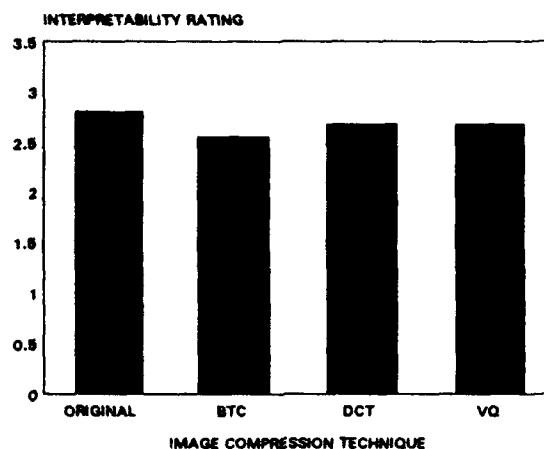


Figure 13. Mean Interpretability Ratings as a Function of Image Compression Method.

Along with the difference found between the compression techniques, a significant difference ($p < 0.01$) was also found between the images used in the study. The individual IA ratings spanned the range of 0 through 8. Most of the IA responses, however, were grouped toward the low end of the rating scale (mode over all images, compression techniques, and IAs was 3.00). The means of the original (i. e., uncompressed) images spanned the range from 1.40 to 5.4.

An attempt was made to explore the relationship between the image quality measurements and the image interpretability ratings. It was expected that image quality (IPR and CR) would be positively correlated with interpretability. That is small IPRs (high resolution) and large CR would be associated with higher ratings. This did not prove to be the case. Regression analyses were performed to determine the exact relationship between the IPR and CR measurements and the IA interpretability rating. The image quality measurements for each image were used to predict the mean IA rating for each compression technique. The r^2 values were below 0.20 in each of the four cases. This does not support the expectation of image quality as a predictor of image interpretability.

Section VII

AUTOMATIC TARGET CUEING

CLUTTER ESTIMATION

Automation is being developed for application to target detection, location, recognition, identification, and aim point selection problems. The goal of these developments is to provide a rapid, accurate, and repeatable capability, overcoming the limitations of the human operator. One of the major difficulties in developing this capability is the generation of false alarms (i. e., the declaration that a target of a specified type is present when the declaration is false). The primary source of false alarms is the presence of target-like signatures within the imaged scene: clutter. Waldman, et al. [22], define image clutter as "the amount of time-independent background texture which is similar to the target in size, shape, and orientation." They further stress that "a scene does not contain some inherent clutter, but only has a clutter value when some target is considered in it." Bhanu [23] describes a "double window filter" employed in the target detection stage of an ATC. He notes that the "inner window surrounds the target, and the outer window contains background." A variation of the double window filter includes a third window, the "guard band," located between the target and background rectangles as shown in Figure 14:

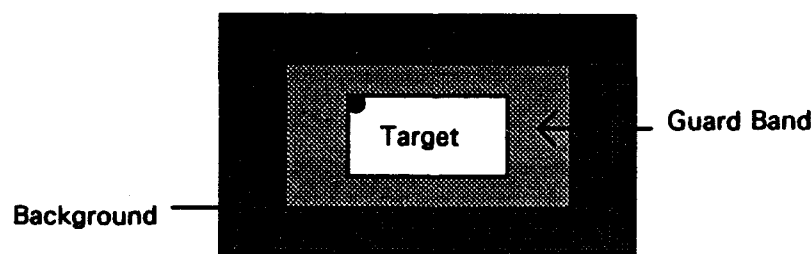


Figure 14. Triple Window Filter.

The size of the target window is selected so as to contain the target of interest, at any orientation, and at the appropriate image scale. The background window region is sized so as to provide a valid estimate of the statistical parameters (e. g., mean and variance) which characterize the scene background in the immediate vicinity of the target. The dimensions of the guard band are selected to minimize confounds between

pixels values arising from the target signature and those that are due solely to background. In the present study, a mobile missile was selected as the target of interest. Image measurements were made from an array of these vehicles positioned at a variety of orientations against a grassy background. The inner (target) rectangle was sized so that it would just overlay any of these signatures. The guard band was sized in both length and height to be equal to one half of the major dimension of the target (in pixels). The background window was dimensioned so as to contain the same number of pixels as the target window.

The triple window filter represents a sampling strategy. If it is positioned at a given pixel location (e. g., the black dot in the upper left corner of the target region in Figure 13), all the other pixels in that vicinity may be categorized as belonging either to the target region or to the background region. Statistical parameters may be estimated from the target and background regions. These parameters may be combined into a metric or test to determine if the current sample is part of the specified target or part of the background. One such metric is the target-to-background interference ratio (TBIR). (The TBIR metric was widely employed as a measure of local target contrast under the Passive Autonomous Infrared Sensor Technology [PAIRSTECH] program [24].) TBIR is defined as:

$$\frac{\bar{I}_T - \bar{I}_B}{\sqrt{\sigma_T \sigma_B}}$$

where \bar{I}_T is the mean of the target window, \bar{I}_B is the mean of the background region, and σ_T and σ_B are the standard deviations of these regions. TBIR values were computed for every pixel location in the target signature array image. The resultant array was iteratively thresholded so as to determine the single TBIR value such that all other TBIR values which were greater to or equal to that value arose only from pixel locations within the targets themselves. That is, no false alarms were produced. The sampling strategy and TBIR metric were then applied all the imagery (original and compressed versions).

The triple window sampling scheme, together with the TBIR metric, serves as a measure of scene clutter. The target window accounts for the size and orientation

attributes. The TBIR comparison is a measure of the separability of a target signature from background. This process intuitively captures the "target-like" definition desired for quantifying scene clutter in terms of the generation of false alarms. When the process is applied to a scene which does not contain the specified target of interest, and the ratio of clutter to total image area is formed, we have a measure of the likelihood that any local area will generate a false alarm.

RESULTS

This process was applied to both the original and compressed versions of the OUE imagery set. Figure 15 graphically depicts the result. As can be seen from the Figure, the DCT and VQ produced elevated clutter values (false alarm candidates) compared to the original images. This elevation was statistically significant ($p < 0.01$) when tested by the paired-comparison t-Test procedure.

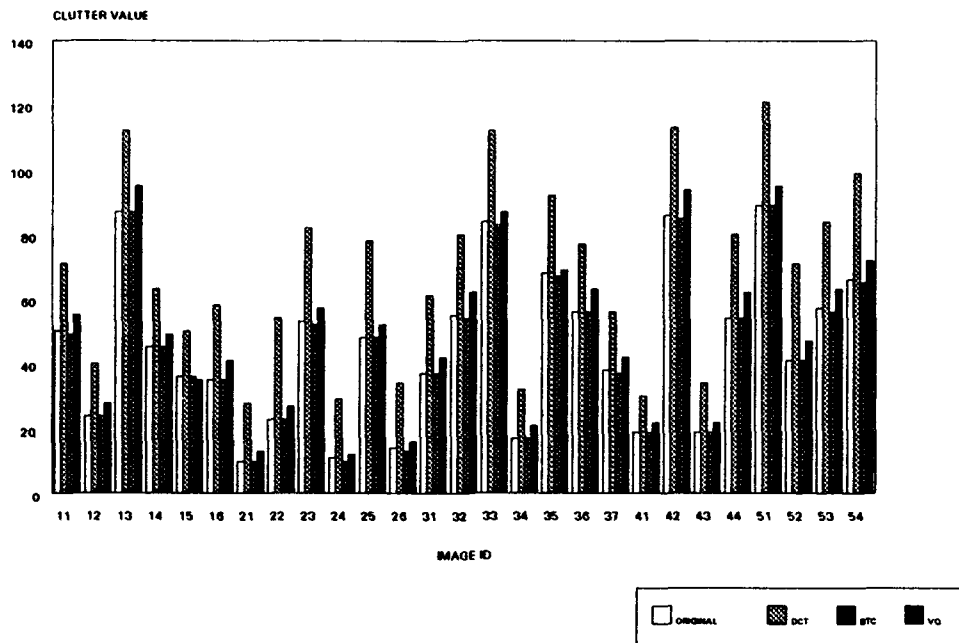


Figure 15. Effects of Image Compression on TBIR-based Clutter Metric.

Section VIII

CONCLUSIONS

This study demonstrated the pronounced effect that different type of image compression algorithm can have on objective measurements of SAR image quality, specifically the IPR. Since IPR has been demonstrated elsewhere to be a determining factor in the ability of the image analyst to extract fine detail information from an image, then any degrading of the SAR IPR can be expected to result in a concomitant loss in image information.

The study (again) demonstrated the weaknesses of image fidelity metrics (MSE and SNR). In general, neither of these measurements correlated significantly with (i. e., served as predictors of) the changes in SAR image quality revealed by the pre- and post-compression image quality measurements. Further, there was relatively poor agreement demonstrated between the two image fidelity measurement techniques used in this study.

A significant difference in operator performance, as predicted by an image interpretability rating scale, was found in comparing the BTC imagery to their original versions. The mean difference was approximately 0.30 interpretability rating scale units. This difference, therefore, may not be operationally significant. The image interpretability rating scale employed in this portion of the study also proved sensitive in identifying differences between individual images.

An attempt was made to demonstrate the predictive validity of the image quality measurements with regard to the image interpretability ratings. This attempt was not successful.

The study revealed a high likelihood that the application of certain image compression techniques may result in an increase in false alarms if an ATR or other automated image exploitation process is applied to compressed imagery. This effect was particularly evident with respect to the DCT technique.

REFERENCES

- [1] Shannon, C.E., "A Mathematical Theory of Communication," Bell Systems Technical Journal, Vol. 27, pp. 379-423, July 1948, pp. 623-656, Aug 1948.
- [2] Frost, V.S., and Minden, G.J., "A Data Compression Technique for Synthetic Aperture Radar Images," IEEE Transactions on Aerospace and Electronic Systems, Vol AES-22, pp. 47-55, Jan 86.
- [3] Wu, C., "Considerations on Data Compression of Synthetic Aperture Radar Images," Proceedings of the SPIE, Vol 87, pp. 134-140, 1976.
- [4] Kashef, B.G., and Tam, K.K., "Synthetic Aperture Radar Image Bandwidth Compression," Proceedings of the SPIE, Vol 432, pp. 45-53, Aug 1983.
- [5] Delp, E., and Mitchell, O., "Image Compression Using Block Truncation Coding," IEEE Transactions on Communications, Vol Com-27, pp. 1335-1342, Sept 1979.
- [6] Linde, Y., Buzo, A., and Gray, R.M., "An Algorithm for Vector Quantization Design," IEEE Transactions on Communications, Vol Com-28, pp. 84-95, January 1980.
- [7] Aravind, R., and Gersho, A., "Low Rate Image Coding with Finite-State Vector Quantization," Proceedings of ICASSP, pp 4.3.1-4.3.3, 1986.
- [8] Gray, R.M., "Vector Quantization," IEEE ASSP Magazine, Vol 1, pp 4-29, Apr 1984.
- [9] Ramamoorthy, P., and Tran, T., "A Hybrid Coding Involving ADM and Vector Quantization for Digital Video Image Compression," IEEE International Conference on Acoustic, Speech and Signal Processing, April 1986.
- [10] Cappellini, V., and Del Re, E., "Image Data Compression by the Discrete Cosine Transform," Mathematics and Computers in Simulation, Vol 27, pp. 599-608, Mar 1985.
- [11] Bracewell, R., *The Fourier Transform and Its Applications*, 2nd edition, McGraw-Hill, New York (1978).

- [12] Makhoul, J., "A Fast Cosine Transform in One and Two Dimensions," IEEE Transactions on Acoustic, Speech and Signal Processing, Vol ASSP-28, pp 27-34, Feb 1980.
- [13] Habibi, A., "A Review of Image Bandwidth Compression Techniques," Proceedings of the SPIE, Vol 528, pp 138-144, Aug 1985.
- [14] Gonzalez, R., and Wintz, P., *Digital Image Processing*, 2nd edition, Addison-Wesley, Reading, Massachusetts (1987).
- [15] Lloyd, S. P., "Least Squares Optimization in PCM," Bell Lab, Tech Note, 1957; also in IEEE Transactions on Information Theory, Vol IT-28, pp 129-137, Mar 1982.
- [16] Max, J., "Quantization for Minimum Distortion," IRE Transactions on Information Theory, Vol IT-6, pp 7-12, Mar 1960.
- [17] Porcello, L., Massey, N., Innes, R., and Marks, J., "Speckle Reduction in Synthetic Aperture Radars," Journal of the Optical Society of America, Vol 66, pp 1305-1311, Nov 1976.
- [18] Ravich, L., "Evolution of Electronic Images," Laser Focus World, Vol 23, pp 145-155, 1987.
- [19] Jain, E., "Image Data Compression: A Review," Proceedings of the IEEE, Vol 69, pp 349-388.
- [20] Hall, E., *Computer Image Processing and Recognition*, Addison-Wesley, Reading, Massachusetts (1979).
- [21] Rosenfeld, A., and Kak, A., *Digital Picture Processing*, 2nd edition, Academic Press, New York (1982).
- [22] Waldman, G., Wootton, J., Hobson, G., and Luetkemeyer, K., "A Normalized Clutter Measure for Images," Computer Vision, Graphics, and Image Processing, **42**, 137-156 (1958).
- [23] Bahnu, B., "Automatic Target Recognition: State of the Art Survey," IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-22, No. 4, July 1986.
- [24] Sajadi, F., and Bazakos, M., "A Perspective on ATR Evaluation Methodology," Proceedings of the International Conference on Performance

Evaluation of Signal and Image Processing Systems, SPIE 1310, April 1990.

APPENDIX

Cross Reference of Study Images with STANAG Target Categories

IMAGE IDENTIFIER

TARGET CATEGORY

11 12 13 14 15 16 21 22 23 24 25 26 31 32 33 34 35 36 37 41 42 43 44

- [illegible]

GLOSSARY

AFB	Air force base
AL	Armstrong Laboratory
ANOVA	Analysis of variance
ATC	Automatic target cueing
ATR	Automatic target recognition
BTC	Block truncation code
C	Signal
C_R	Compression ratio
CR	Contrast ratio
D	Dimension
dB	Decibel
DCT	Discrete cosine transform
DoD	Department of Defense
DPCM	Differential pulse code modulation
FASTC	Foreign Aerospace Science and Technology Center
FFT	Fast Fourier transform
I	In-phase
IA	Image analyst
IPR	Impulse response

ISL	Integrated side lobe
JPEG	Joint photographic exploitation group
LUT	Look-up-table
MSE	Mean square error
N_a	Additive noise
N_r	Receiver thermal noise
OUE	Operational utility evaluation
P_{CC}	Pearson correlation coefficient
Pixel	Picture element
Q	Quadrature-phase
SAR	Synthetic aperture radar
SP	Self-propelled
S_n	System noise
SNR	Signal-to-noise ratio
STANAG	Standardization agreement
TBIR	Target-to-background interference ratio
USMC	United States Marine Corps
VQ	Vector quantization